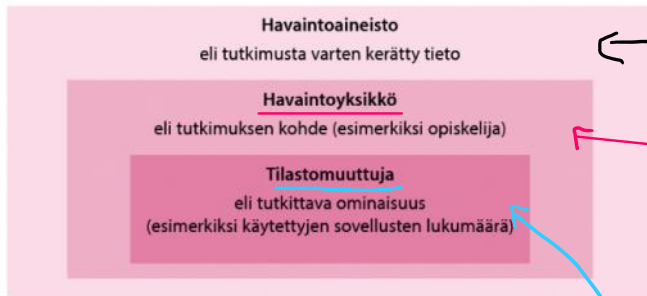


Tutkimusaineisto ja sen kerääminen



← perusjoukko / populaatio

kokonaismaineisto
(kaikki mukana) tai otos
(rajattu joukko)

diskreetti (yksittäinen arvo)
tai
jatkuva (pyöristetty arvo)

Esim. Tutkitaan, mikä on MAB5 kurssin tavoitearvosana.

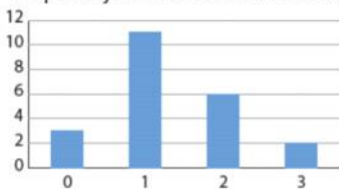
7, 8, 8, 8, 8, 9, 10

Frekvenssi
Frekvenssi f ilmaisee, kuinka monta kertaa muuttujan arvo esiintyy havaintoaineistossa. Suhteellinen frekvenssi $f\%$ ilmaisee, kuinka suuri prosenttiosuus muuttujan arvolla on kaikista havainnoista.

muuttuja arvosana	frekvenssi f	$f\%$
7	1	$\frac{1}{7} = 0,142... \approx 14\%$
8	4	$\frac{4}{7} \approx 57\%$
9	1	14%
10	1	14%
	yht. 7	

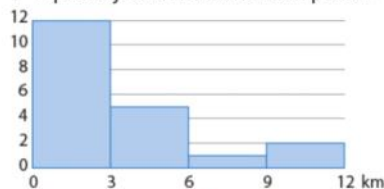
- Pylväsdiagrammilla tai histogrammilla** havainnollistetaan frekvenssi-jakaumia. Pylväiden korkeus määräytyy frekvenssin mukaan.

f Opiskelijoiden sisarusten lukumäärä



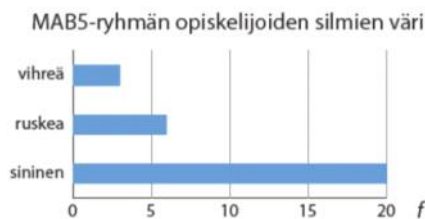
Diskreetin muuttujan jakaumaa kuvataan pylväsdiagrammilla, jossa pylväät on piirretty erilleen.

f Opiskelijoiden koulumatkan pituus

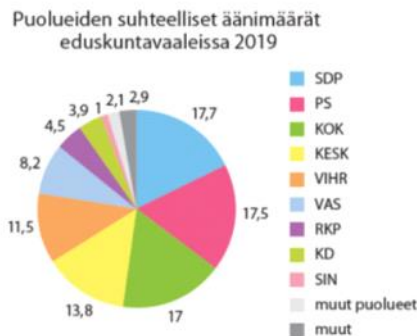


Jatkuvan muuttujan jakaumaa kuvataan histogrammilla, jossa pylväät on piirretty yhteen.

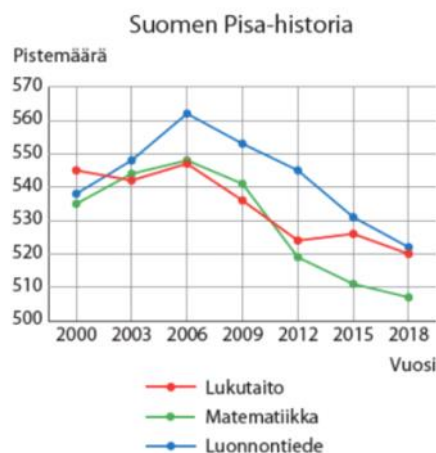
- **Vaakapylväitä** käytetään silloin kun muuttujan arvoilla ei ole luontevaa suuruusjärjestystä.



- **Sektoridiagrammilla** havainnollistetaan kokonaisuuden jakaantumista osiin. Sektorin keskuskulman suuruus määräytyy suhteellisen frekvenssin mukaan.



- **Viivadiagrammia** käytetään erityisesti silloin, kun halutaan havainnollistaa jonkin asian kehittymistä tai vaihtelua ajan kuluessa.



Kun samasta tutkittavasta asiasta on havaintoja peräkkäisten, yhtä pitkien ajanjaksojen ajan, tilastoa kutsutaan **aikasarjaksi**.

1.3 Opettaja seurasi viikon ajan, kuinka monta kertaa erään ryhmän opiskelijat myöhästyivät aamun ensimmäiseltä tunnilta. Tulokset on esitetty oheisessa taulukossa.

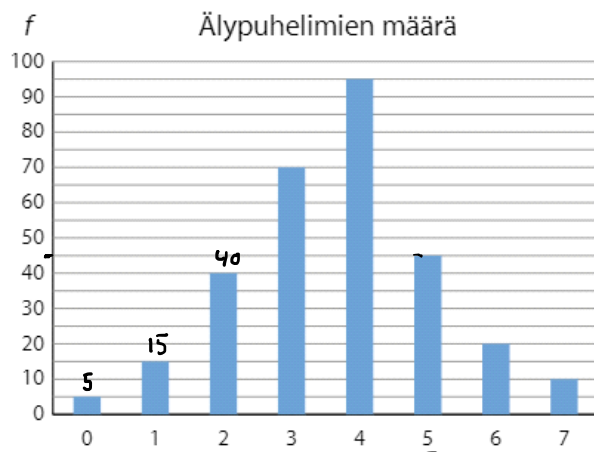
- Kuinka moni opiskelija ei myöhästynyt kertaakaan? **5**
- Kuinka moni opiskelija myöhästyi vähintään kahtena aamuna? **12**
- Kuinka monta opiskelijaa ryhmässä oli? **25**
- Laske muuttujan arvon 2 suhteellinen frekvenssi.
Mitä tämä kuvaa? **prosenttiosuus aineistossa**

Myöhästymisten lukumäärä	f
0	5
1	8
2	6
3	3
4	2
5	1

suhteellinen frekvenssi
 $f\%$

$$\frac{6}{25} = 0,24 = 24\%$$

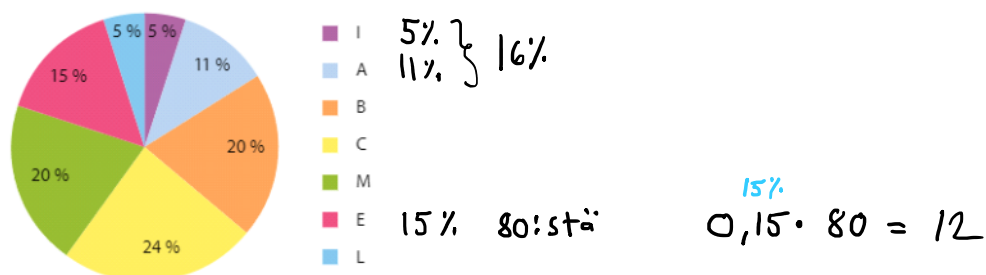
1.4 Tutkimuksessa selvitettiin kotitalouksissa olevien älypuhelimien määriä.



- a) Kuinka monessa kotitaloudessa on viisi älypuhelimia? **45**
- b) Kuinka monessa kotitaloudessa on korkeintaan kaksi älypuhelimia? **60**
- c) Kuinka monta kotitaloutta osallistui tutkimukseen? $5+15+40+70+95+45+20+10 = 300$

1.6 Ylioppilastutkintolautakunta päättää eri aineiden arvosanojen piste-rajat kullakin tutkintokerralla erikseen. Koulun lyhyen matematiikan ylioppilaskokeen arvosanat jakaantuivat diagrammin mukaisesti.

Ylioppilaskokeen arvosanat



- a) Kaksi heikointa arvosanaa ovat I ja A. Kuinka monta prosenttia kokeeseen osallistujista sai jommankumman heikoimmista arvosanoista?
- b) Koulussa lyhyen matematiikan kirjoitti 80 opiskelijaa. Kuinka moni sai arvosanan E?

1.16 Taulukkoon on koottu erään koulun

A-englannin yo-kokeen arvosanjakauma
pistemäärien avulla ilmoitettuna ($L = 7$,
 $E = 6$, $M = 5$, $C = 4$, $B = 3$, $A = 2$ ja $I = 0$).

- a) Mikä on aineiston muuttuja? pistemäärä
- b) Onko muuttuja jatkuva vai diskreetti?
- c) Mikä on yleisin arvosana?
- d) Kuinka monta opiskelijaa kyseisessä koulussa osallistui kokeeseen? 118
- e) Kuinka monta prosenttia opiskelijoista läpäisi A-englannin yo-kokeen, kun arvosana I on hylätty?

$$118 - 6 = 112$$

$$112/118$$

$$= 0,94915254237288135593 \approx 95\%$$

Pistemäärä	f
0	6
2	13
3	24
<u>4</u>	<u>29</u>
5	23
6	18
7	5

$$6+13+24+29+23+18+5 \\ = 118$$